Stochastic monotonicity in Young graph and Thoma theorem.

Alexey Bufetov[*] Vadim Gorin[†]

May 22, 2018

Abstract

We show that the order on probability measures, inherited from the dominance order on the Young diagrams, is preserved under natural maps reducing the number of boxes in a diagram by 1. As a corollary we give a new proof of the Thoma theorem on the structure of characters of the infinite symmetric group.

We present several conjectures generalizing our result. One of them (if it is true) would imply the Kerov's conjecture on the classification of all homomorphisms from the algebra of symmetric functions into $\mathbb{R}$ which are non-negative on Hall–Littlewood polynomials.

1 Introduction

1.1 Problem setup and results

For a number $n = 0, 1, 2, \dots$, a partition λ of n is a sequence of integers $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ such that $|\lambda| = n$, where $|\lambda| = \sum_{i=1}^{\infty} \lambda_i$. We identify a partition λ with the Young diagram, which is a collection of $|\lambda|$ boxes with positive coordinates (i, j) forming the following set

$$\{(i, j) \subset \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \mid j \leq \lambda_i\}.$$

When drawing pictures we adopt the notation that the first index i increases as we move down, while the second index j increases as we move to the right, cf. Figures 1 and 2.

The Young graph $\mathbb{Y} = \bigcup_{n=0}^{\infty} \mathbb{Y}_n$ is a graded graph such that the vertices of $\mathbb{Y}_n$ are all partitions of n. In particular, $\mathbb{Y}_0$ contains only the empty partition $\emptyset = (0, 0, \dots)$. An

[*]International Laboratory of Representation Theory and Mathematical Physics, Department of Mathematics, Higher School of Economics, Moscow, Russia, and Institute for Information Transmission Problems of Russian Academy of Sciences, Moscow, Russia. E-mail: alexey.bufetov@gmail.com

[†]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA, and Institute for Information Transmission Problems of Russian Academy of Sciences, Moscow, Russia. E-mail: vadicgor@gmail.com

edge joins $\lambda \in \mathbb{Y}_n$ with $\mu \in \mathbb{Y}_{n-1}$, $n \geq 1$, if and only if λ differs from μ by the addition of a single box, which we denote $\mu \nearrow \lambda$.

For a Young diagram λ, its *dimension*[1] denoted by $\dim(\lambda)$ is the number of oriented paths in $\mathbb{Y}$ which start at $\emptyset$ and end at λ.

Let M_n be a probability measure on $\mathbb{Y}_n$. Its *projection* onto $\mathbb{Y}_{n-1}$ denoted by $\pi^n_{n-1} M_n$ is defined via

$$(\pi^n_{n-1} M_n)(\mu) = \sum_{\lambda \in \mathbb{Y}_n :\, \mu \nearrow \lambda} \frac{\dim(\mu)}{\dim(\lambda)} M_n(\lambda).$$

The definition readily implies that $\pi^n_{n-1} M_n$ is a probability measure. Iterating the maps $M_n \mapsto \pi^n_{n-1} M_n$ one similarly defines the projection of M_n onto M_k, $0 \leq k < n$, denoted by $\pi^n_k M_n$.

Definition 1.1. *A sequence of measures $\{M_n\}_{n=0}^{\infty}$ is called a* coherent system *on $\mathbb{Y}$ if each M_n, $n = 0, 1, \dots$ is a probability measure on $\mathbb{Y}_n$ and for any $0 \leq k < n$ the measure M_k is the projection of M_n onto $\mathbb{Y}_k$, i.e. $M_k = \pi^n_k M_n$.*

In last 40 years coherent systems on $\mathbb{Y}$ were enjoying lots of interest due to their connections to several seemingly unrelated topics. First, one can show that they are in bijection with normalized characters for the infinite symmetric group and have a close relation to the finite factor and spherical representations of the latter, see [VK1], [K2], [Ok1]. Second, there is a correspondence between such systems of measures and totally positive upper triangular Toeplitz matrices, see [T], [K2, Section 2.2], [Ok1]. Third, they are naturally linked to combinatorial objects appearing in the study of the Robinson–Schensted–Knuth correspondence, cf. [VK2]. Finally, several instances of these systems, e.g. the celebrated Plancherel distributions, exhibit a remarkable asymptotic behavior as $n \to \infty$ and, in particular, numerous connections to random matrices, see [BDJ], [BOO], [Ok1], [J], [K1], [IO].

The classification of all coherent systems on $\mathbb{Y}$ (in an equivalent form) is now known as *Thoma theorem*. Its formulation uses the symmetric functions notations which we now introduce. Let Λ be the algebra of all symmetric functions in countably many variables $x_1, x_2, \dots$, see e.g. [Ma, Chapter 1, Section 2]. One way to define Λ is as an algebra (over $\mathbb{R}$) of polynomials in Newton power sums p_k, $k = 1, 2, \dots$

$$p_k = x_1^k + x_2^k + x_3^k + \dots .$$

An important linear basis of Λ is formed by *Schur symmetric functions* s_λ, $\lambda \in \mathbb{Y}$, and we refer to [Ma, Chapter 1, Section 3] for the exact definition and properties of s_λ.

We also define Ω to be the set of all pairs of sequences $(\alpha, \beta) = (\alpha_1 \geq \alpha_2 \geq \dots \geq 0, \beta_1 \geq \beta_2 \geq \dots \geq 0)$, such that $\sum_{i=1}^{\infty} (\alpha_i + \beta_i) \leq 1$.

Theorem 1.2 (Thoma theorem, cf. [T], [VK1], [Ok1], [KOO], [V2])**.** *The set of all coherent systems is a (Choquet) simplex, whose extreme points are parameterized by elements*

[1] The name originates in the fact that $\dim(\lambda)$ coincides with the dimension of the irreducible representation of the symmetric group $\mathfrak{S}_n$ indexed by λ. Here $n = |\lambda|$.

of Ω. The extreme system of measures $\{M_n^{(\alpha,\beta)}\}_{n=0}^{\infty}$ parameterized by $(\alpha,\beta)\in\Omega$ is given by

$$M_n^{(\alpha,\beta)}(\lambda)=\dim(\lambda)s_\lambda(\alpha,\beta), \tag{1}$$

where $s_\lambda(\alpha,\beta)$ is the image of s_λ under the algebra homomorphism from Λ to $\mathbb{R}$ defined on power sums p_k via

$$p_1\mapsto p_1(\alpha,\beta)=1,\quad p_k\mapsto p_k(\alpha,\beta)=\sum_{i=1}^{\infty}\alpha_i^k+(-1)^{k-1}\sum_{i=1}^{\infty}\beta_i^k,\ k=1,2,\dots. \tag{2}$$

One of the aims of our article is to give a new proof of Theorem 1.2 based on a monotonicity–preservation property that we will now present. Our proof of Thoma theorem is based on the combinatorial and probabilistic ideas only; other existing proofs use highly nontrivial analytic [T] or algebraic [VK1, KOO, Ok1] methods (see, however, [V2]). We hope that the strategy used in our proof of Theorem 1.2 could be used in the future to establish the validity of a generalization of Theorem 1.2 known as the *Kerov's conjecture*, see Section 1.2 for more details.

Let us equip $\mathbb{Y}_n$ with a partial order known as *dominance order*. For $\lambda,\mu\in\mathbb{Y}_n$ we write $\lambda\geq\mu$, if for all $k=1,2,\dots$ we have

$$\lambda_1+\lambda_2+\cdots+\lambda_k\geq\mu_1+\mu_2+\cdots+\mu_k.$$

Further, we say that a measure ρ on $\mathbb{Y}_n$ is an atom if its support consists of a single element and write $\sup(\rho)$ for this element. Note that we allow the mass of ρ to be different from 1 here.

Definition 1.3. *Let ρ and ρ' be two measures on $\mathbb{Y}_n$ of the same total mass, i.e. $\rho(\mathbb{Y}_n)=\rho'(\mathbb{Y}_n)$. We say that ρ stochastically dominates ρ' and write $\rho\geq\rho'$, if there exist $k>0$ and $2k$ measures $\rho_1,\dots,\rho_k$, $\rho'_1,\dots,\rho'_k$, such that $\rho=\sum_{i=1}^k\rho_i$, $\rho'_i=\sum_{i=1}^k\rho'_i$, and, moreover, ρ_i, ρ'_i are atoms of the same mass and with $\sup(\rho_i)\geq\sup(\rho'_i)$ for each $i=1,\dots,k$.*

Informally, Definition 1.3 means that ρ can be obtained from ρ' by moving masses up with respect to our partial order.

Theorem 1.4. *Take $0\leq k<n$ and let ρ and ρ' be two measures on $\mathbb{Y}_n$ of the same total mass. If $\rho\geq\rho'$, then the same is true for their projections on $\mathbb{Y}_k$, i.e. $\pi_k^n\rho\geq\pi_k^n\rho'$.*

We prove Theorem 1.4 in Section 2. Our proof is based on inequalities for the dimensions in Young graph presented in Corollary 2.6. We also explain that these inequalities admit natural generalizations to the statements about the monomial positivity of certain quadratic expressions in Schur polynomials; we do not know a proof for the latter monomial positivity and present it as Conjecture 2.2.

In Section 4, we combine Theorem 1.4 with the Law of Large Numbers for a subclass of extreme coherent systems and deduce Theorem 1.2. Finally, in Section 3 we recall the aforementioned Law of Large Numbers and explain several known strategies of its proof.

1.2 t–Deformation and Kerov's conjecture

Theorem 1.2 is known (see e.g. [K2]) to be equivalent to the following description of all Schur–positive homomorphisms from Λ into $\mathbb{R}$.

Theorem 1.5. *The set of algebra homomorphisms $\varrho : \Lambda \to \mathbb{R}$ normalized by the condition $\varrho(p_1) = 1$ and such that $\varrho(s_\lambda) \geq 0$ for all $\lambda \in \mathbb{Y}$, is in bijection with Ω. The homomorphism corresponding to $(\alpha, \beta) \in \Omega$ is defined by its values on power sums p_k*

$$p_1 \mapsto p_1(\alpha, \beta) = 1, \quad p_k \mapsto p_k(\alpha, \beta) = \sum_{i=1}^{\infty} \alpha_i^k + (-1)^{k-1} \sum_{i=1}^{\infty} \beta_i^k, \ k = 1, 2, \dots. \tag{3}$$

A natural way to generalize Theorem 1.5 is by replacing Schur functions s_λ by other classes of symmetric functions. Kerov conjectured 20 years ago that when s_λ are replaced by their celebrated (q,t)-deformation — Macdonald polynomials $M_\lambda(\cdot; q, t)$ — then (for $0 \leq q < 1$, $0 \leq t < 1$) the Macdonald–positive homomorphisms are still in bijection with elements of Ω. The conjectural correspondence is established through the formulas very similar to (3), see [K2, Chapter II, Section 9] for the details. The completeness of the Kerov's list of homomorphisms is still an open problem (though it is relatively easy to show that all these homomorphisms are indeed Macdonald–positive, see e.g. [BC, Section 2.2.1]). Recently, these homomorphisms have been actively used for the asymptotic analysis of a variety of probabilistic systems in the framework of Macdonald processes, see [BC], [BCGS].

The $q = 0$ versions of Macdonald polynomials are the Hall–Littlewood polynomials, see [Ma]. This particular case of the Kerov's conjecture is especially interesting, since when $t = p^{-1}$ the conjecture is equivalent to the (conjectural) classification of all conjugation invariant ergodic measures on infinite uni–uppertriangular matrices over a finite field with p elements $\mathbb{F}_p$, see [GKV, Section 4].

Recently a progress on the t–deformation of Theorem 1.2 (equivalent to the Hall-Littlewood case of Kerov's conjecture, see [GKV, Section 4] and [Fu, Section 4.2] for the details) was achieved in [BP], where the Law of Large Numbers for the measures arising in it was proved. We thus hope that our approach to the proof of Theorem 1.2 can be extended to the Hall–Littlewood case of Kerov's conjecture. More precisely, if one tries to mimic our approach, then the conjecture at $t = p^{-1}$ reduces to the following inequality.

Let U_n be the group of all uni–uppertriangular matrices over $\mathbb{F}_p$. Note that for each $u \in U_n$ all its eigenvalues are 1s and thus we can assign to it a unique Young diagram $\mathcal{J}(u) \in \mathbb{Y}_n$ whose row lengths are sizes of the blocks in Jordan Normal Form of u. We define

$$\dim_t(\lambda) = |\{u \in U_n \mid \mathcal{J}(u) = \lambda\}|.$$

Further, for any $u \in U_n$ we set $u^{(n-1)} \in U_{n-1}$ to be its top–left $(n-1) \times (n-1)$ corner, and define for $\mu \in \mathbb{Y}_{n-1}$, $\lambda \in \mathbb{Y}_n$

$$\dim_t(\mu \nearrow \lambda) = |\{u \in U_n \mid \mathcal{J}(u^{(n-1)}) = \mu,\ \mathcal{J}(u) = \lambda\}|.$$

We remark that [B, Theorem 2.3] (see also [Kir]) gives an explicit formula for the ratio $\frac{\dim_t(\mu \nearrow \lambda)}{\dim_t(\mu)}$, which, in particular, implies that $\dim_t(\mu \nearrow \lambda)$ vanishes unless $\mu \nearrow \lambda$.

Conjecture 1.6. *Let $\lambda, \hat{\lambda} \in Y_n$ and $\mu, \hat{\mu} \in Y_{n-1}$ be two pairs of Young diagrams, such that both λ,$\hat{\lambda}$ and μ,$\hat{\mu}$ differ by the move of box (i,j) into the position $(\hat{i},\hat{j})$ with $\hat{i} > i$. Further, assume that $\lambda \setminus \mu = \hat{\lambda} \setminus \hat{\mu} = (r,c)$, cf. Figures 1 and 2. If $r < i$ then*

$$\frac{\dim_t(\hat{\mu} \nearrow \hat{\lambda})}{\dim_t(\hat{\lambda})} \geq \frac{\dim_t(\mu \nearrow \lambda)}{\dim_t(\lambda)}. \tag{4}$$

If $r > \hat{i}$, then

$$\frac{\dim_t(\hat{\mu} \nearrow \hat{\lambda})}{\dim_t(\hat{\lambda})} \leq \frac{\dim_t(\mu \nearrow \lambda)}{\dim_t(\lambda)}. \tag{5}$$

This conjecture can be also restated as a certain inequality for the values of Hall–Littlewood polynomials, and its generalization is formulated below in Conjecture 2.4. Computer checks supply the validity of these conjectures, but we have not found a proof.

At $t = 1$ the Hall–Littlewood polynomials turn into the monomial symmetric functions and this case of the Kerov's conjecture is equivalent to the Kingman's classification theorem for exchangeable partition structures on $\mathbb{Z}_{>0}$, see [K2, Chapter I]. Both ingredients of our approach, which are the $t = 1$ versions of Conjecture 1.6 and the Law of Large Numbers for the extreme coherent systems are especially simple and transparent in this case. Thus, by mimicking our proof of Theorem 1.2 one can also get a new proof of the Kingman's classification theorem [Kin].

Acknowledgements. A. B. was partially supported by Simons Foundation-IUM scholarship, by "Dynasty" foundation, and by the RFBR grant 13-01-12449. V. G. was partially supported by the NSF grant DMS-1407562.

2 Monotonicity in Young graph

This section is devoted to the proof of Theorem 1.2.

2.1 Elementary moves

First, let us introduce several additional notations. We say that two distinct Young diagrams $\lambda \in \mathbb{Y}_n$ and $\hat{\lambda} \in \mathbb{Y}_n$ differ by the move of box (i,j) into the position $(\hat{i},\hat{j})$, if there exists $\mu \in \mathbb{Y}_{n-1}$ such that $\mu = \lambda \setminus (i,j) = \hat{\lambda} \setminus (\hat{i},\hat{j})$, see Figure 1 for an illustration. Note that we should have $\hat{i} \neq i$ Further, if $\hat{i} > i$, then $\lambda \geq \hat{\lambda}$ and if $\hat{i} < i$, then $\lambda \leq \hat{\lambda}$.

Recall that for a Young diagram λ, the numbers $\lambda'_1 \geq \lambda'_2 \geq \ldots$ are defined as the column lengths of λ, formally

$$\lambda'_j = |\{i \in \mathbb{Z}_{>0} : \lambda_i \geq j\}|.$$

We also set $\ell(\lambda)$ to be the number of non-zero rows in λ, i.e. $\ell(\lambda) = \lambda'_1$.

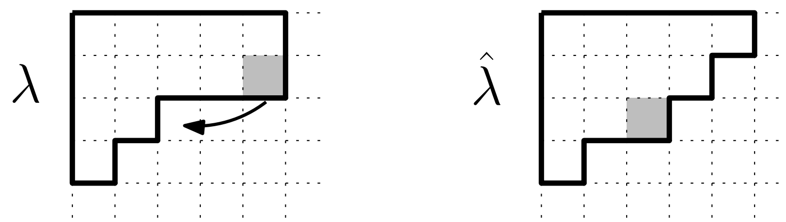

Figure 1: Young diagrams λ and $\hat{\lambda}$ differing by the move of the box $(2,5)$ into the position $(3,3)$. Here $\lambda \geq \hat{\lambda}$ and also $\lambda \succ \hat{\lambda}$.

We evoke the (N–variable version of) Schur symmetric function s_λ. For any $N = 1, 2, \dots$ and Young diagram $\lambda \in \mathbb{Y}$ such that $\ell(\lambda) \leq N$, we have

$$s_\lambda(x_1, \dots, x_N) = \frac{\det_{i,j=1}^{N} \left[x_i^{\lambda_j + N - j} \right]}{\prod_{1 \leq i < j \leq N} (x_i - x_j)}.$$

Finally, we use the notation 1^N for $(\underbrace{1, \dots, 1}_{N})$.

Our proof of Theorem 1.4 relies on the following statement.

Proposition 2.1. *Let $\lambda, \hat{\lambda} \in Y_n$ and $\mu, \hat{\mu} \in Y_{n-1}$ be two pairs of Young diagrams, such that both λ,$\hat{\lambda}$ and μ,$\hat{\mu}$ differ by the move of box (i,j) into the position $(\hat{i},\hat{j})$ with $\hat{i} > i$. Further, assume that $\lambda \setminus \mu = \hat{\lambda} \setminus \hat{\mu} = (r,c)$, cf. Figure 2. Fix any integer $N \geq \ell(\hat{\lambda})$. If $r < i$ then*

$$s_\lambda(1^N) s_{\hat{\mu}}(1^N) \geq s_{\hat{\lambda}}(1^N) s_\mu(1^N). \tag{6}$$

If $r > \hat{i}$, then

$$s_\lambda(1^N) s_{\hat{\mu}}(1^N) \leq s_{\hat{\lambda}}(1^N) s_\mu(1^N). \tag{7}$$

Proof. We recall the *Weyl dimension formula* (see e.g. [Ma, Section 3, Exerceise 1])

$$s_\lambda(1^N) = \prod_{1 \leq a < b \leq N} \frac{\lambda_a - a - \lambda_b + b}{b - a}$$

and plug it into (6). Since $\lambda_a = \hat{\lambda}_a$ and $\mu_a = \hat{\mu}_a$ for $a \neq i, \hat{i}$, many factors on the left and right side cancel out, and (6) turns into

$$\prod_{\substack{1 \leq a \leq N: \\ a \neq i}} |\lambda_a - a - \lambda_i + i| \prod_{\substack{1 \leq a \leq N: \\ a \neq \hat{i}}} |\lambda_a - a - \lambda_{\hat{i}} + \hat{i}| \prod_{\substack{1 \leq a \leq N: \\ a \neq i}} |\hat{\mu}_a - a - \hat{\mu}_i + i| \prod_{\substack{1 \leq a \leq N: \\ a \neq \hat{i}}} |\hat{\mu}_a - a - \hat{\mu}_{\hat{i}} + \hat{i}|$$
$$\stackrel{?}{\geq} \prod_{\substack{1 \leq a \leq N: \\ a \neq i}} |\hat{\lambda}_a - a - \hat{\lambda}_i + i| \prod_{\substack{1 \leq a \leq N: \\ a \neq \hat{i}}} |\hat{\lambda}_a - a - \hat{\lambda}_{\hat{i}} + \hat{i}| \prod_{\substack{1 \leq a \leq N: \\ a \neq i}} |\mu_a - a - \mu_i + i| \prod_{\substack{1 \leq a \leq N: \\ a \neq \hat{i}}} |\mu_a - a - \mu_{\hat{i}} + \hat{i}|$$

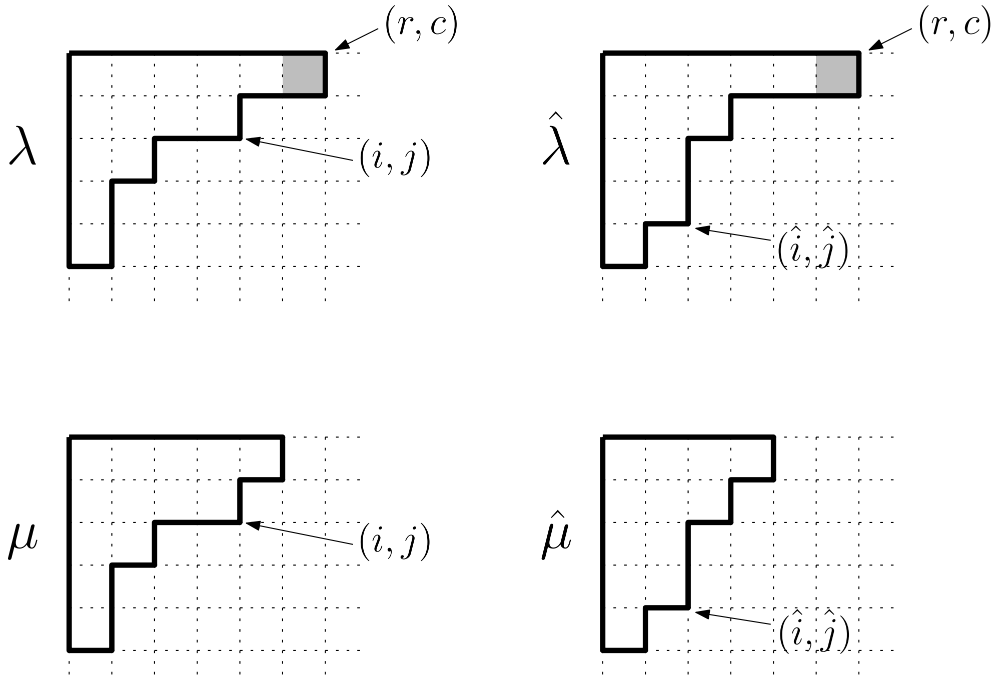

Figure 2: An example of Young diagrams $\lambda, \hat{\lambda}$ and $\mu, \hat{\mu}$ as in Proposition 2.1. Here the gray box is $(r,c) = (1,6)$, and $(i,j) = (2,4)$, $(\hat\imath, \hat\jmath) = (4,2)$.

Since $\lambda_a = \mu_a$ and $\hat\lambda_a = \hat\mu_a$ for $a \neq r$, we can further cancel out the factors to get

$$(\lambda_r - r - \lambda_i + i)(\lambda_r - r - \lambda_{\hat\imath} + \hat\imath)(\hat\mu_r - r - \hat\mu_i + i)(\hat\mu_r - r - \hat\mu_{\hat\imath} + \hat\imath)$$
$$\overset{?}{\geq} (\hat\lambda_r - r - \hat\lambda_i + i)(\hat\lambda_r - r - \hat\lambda_{\hat\imath} + \hat\imath)(\mu_r - r - \mu_i + i)(\mu_r - r - \mu_{\hat\imath} + \hat\imath).$$

Rewriting everything in terms of the parts of λ, we get an equivalent inequality

$$(\lambda_r - r - \lambda_i + i)(\lambda_r - r - \lambda_{\hat\imath} + \hat\imath)(\lambda_r - r - \lambda_i + i)(\lambda_r - r - \lambda_{\hat\imath} + \hat\imath - 2)$$
$$\overset{?}{\geq} (\lambda_r - r - \lambda_i + i + 1)(\lambda_r - r - \lambda_{\hat\imath} + \hat\imath - 1)(\lambda_r - r - \lambda_i + i - 1)(\lambda_r - r - \lambda_{\hat\imath} + \hat\imath - 1).$$

Further transforming, and denoting $\lambda_r - r - \lambda_i + i = x$, $\lambda_r - r - \lambda_{\hat\imath} + \hat\imath - 1 = y$, we get

$$x^2(y^2-1) \overset{?}{\geq} (x^2-1)y^2. \tag{8}$$

Now when $r < i < \hat\imath$, then $y \geq x > 0$ and (8) holds. Similarly, when $r > \hat\imath > i$, then $0 > y \geq x$ and the inequality opposite to (8) holds. □

Based on computer computations we believe that the following two generalizations of Proposition 2.1 should hold.

Recall that a symmetric function $f(x_1, x_2, \dots)$ is called *monomial positive* if the coefficients of its expansion into monomials are non-negative.

Conjecture 2.2. *Let $\lambda, \hat{\lambda} \in Y_n$ and $\mu, \hat{\mu} \in Y_{n-1}$ be two pairs of Young diagrams, such that both λ,$\hat{\lambda}$ and μ,$\hat{\mu}$ differ by the move of box (i,j) into the position $(\hat{\imath},\hat{\jmath})$ with $\hat{\imath} > i$. Further, assume that $\lambda \setminus \mu = \hat{\lambda} \setminus \hat{\mu} = (r,c)$, cf. Figure 2. If $r < i$ then $s_\lambda s_{\hat{\mu}} - s_{\hat{\lambda}} s_\mu$ is monomial–positive. If $r > \hat{\imath}$, then $s_{\hat{\lambda}} s_\mu - s_\lambda s_{\hat{\mu}}$ is monomial–positive.*

Remark 2.3. Monomial positivity (and even stronger Schur–positivity) of similar quadratic expressions has been intensively studied, see [LPP], [LP] and references therein. However it seems that the differences of the form $s_\lambda s_{\hat{\mu}} - s_{\hat{\lambda}} s_\mu$ are out of the scope of those articles.

Further, we recall the definition of (N–variable version of) Hall–Litlewood symmetric function on a parameter $t \in \mathbb{R}$, and a Young diagram λ such that $\ell(\lambda) \le N$, cf. [Ma, Chapter III]

$$Q_\lambda(x_1, \dots, x_N; t) = (1-t)^N \prod_{i=1}^{N-\ell(\lambda)} \frac{1}{1-t^i} \cdot \sum_{\sigma \in \mathfrak{S}(n)} x_{\sigma(1)}^{\lambda_1} \cdots x_{\sigma(N)}^{\lambda_N} \prod_{1 \le i < j \le N} \frac{x_{\sigma(i)} - t x_{\sigma(j)}}{x_{\sigma(i)} - x_{\sigma(j)}}.$$

Note the normalization that we use, and which is the same as in [Ma].

Conjecture 2.4. *Suppose that $0 \le t \le 1$ and let $\lambda, \hat{\lambda} \in Y_n$ and $\mu, \hat{\mu} \in Y_{n-1}$ be two pairs of Young diagrams, such that both λ,$\hat{\lambda}$ and μ,$\hat{\mu}$ differ by the move of box (i,j) into the position $(\hat{\imath},\hat{\jmath})$ with $\hat{\imath} > i$. Further, assume that $\lambda \setminus \mu = \hat{\lambda} \setminus \hat{\mu} = (r,c)$, cf. Figure 2. Fix any integer $N \ge \ell(\hat{\lambda})$. If $r < i$ then*

$$\left(1 - t^{\hat{\lambda}'_c - \hat{\lambda}'_{c+1}}\right) \frac{Q_{\hat{\mu}}(1^N; t)}{Q_{\hat{\lambda}}(1^N; t)} \ge \left(1 - t^{\lambda'_c - \lambda'_{c+1}}\right) \frac{Q_\mu(1^N; t)}{Q_\lambda(1^N; t)}. \tag{9}$$

If $r > \hat{\imath}$, then

$$\left(1 - t^{\hat{\lambda}'_c - \hat{\lambda}'_{c+1}}\right) \frac{Q_{\hat{\mu}}(1^N; t)}{Q_{\hat{\lambda}}(1^N; t)} \le \left(1 - t^{\lambda'_c - \lambda'_{c+1}}\right) \frac{Q_\mu(1^N; t)}{Q_\lambda(1^N; t)}. \tag{10}$$

Remark 2.5. When $t = 0$, Conjecture 2.4 turns into Proposition 2.1. When $t = 1$, the Hall–Littlewood functions $Q_\lambda(\cdot; t)$ turn into the monomial symmetric functions and the validity of Conjecture 2.4 can be similarly established (in fact, inequalities turn into equalities in this case). For general t we are not aware of any simple analogues of the Weyl dimension formula for $P_\lambda(1^N; t)$ and the strategy employed in the proof of Proposition 2.1 fails.

2.2 Proof of Theorem 1.4

The following statement is an immediate corollary of Proposition 2.1.

Corollary 2.6. *Let $\lambda, \hat{\lambda} \in Y_n$ and $\mu, \hat{\mu} \in Y_{n-1}$ be two pairs of Young diagrams, such that both λ,$\hat{\lambda}$ and μ,$\hat{\mu}$ differ by the move of box (i,j) into the position $(\hat{\imath},\hat{\jmath})$ with $\hat{\imath} > i$. Further, assume that $\lambda \setminus \mu = \hat{\lambda} \setminus \hat{\mu} = (r,c)$, cf. Figure 2. If $r < i$, then*

$$\frac{\dim(\hat{\mu})}{\dim(\hat{\lambda})} \ge \frac{\dim(\mu)}{\dim(\lambda)}. \tag{11}$$

If $r > \hat{i}$*, then*

$$\frac{\dim(\hat{\mu})}{\dim(\hat{\lambda})} \le \frac{\dim(\mu)}{\dim(\lambda)}. \tag{12}$$

Proof. The statement follows from Proposition 2.1 and the limit relation

$$\dim(\lambda) = \lim_{N\to\infty} \frac{s_\lambda(1^N)}{N^{|\lambda|}}.$$

The simplest way to prove the latter limit identity is through the explicit formulas for $\dim(\lambda)$ and $s_\lambda(1^N)$, see e.g. [Ma, Chapter I, Section 3, Examples 4-5 and Section 5, Example 2].

Alternatively, one can directly prove (11), (12) along the lines of the proof of Proposition 2.1. □

Remark 2.7. Conjecture 1.6 can be obtained from Conjecture 2.4 in the same way as Corollary 2.6 follows from Proposition 2.1.

Definition 2.8. *For two Young diagrams* $\lambda, \hat{\lambda} \in \mathbb{Y}_n$ *we say that* λ *covers* $\hat{\lambda}$ *and write* $\lambda \succ \hat{\lambda}$ *if* λ *and* $\hat{\lambda}$ *differ by the move of the box* $(i,j) \subset \lambda$ *into the position* $(\hat{i},\hat{j}) \subset \hat{\lambda}$ *such that either* $\hat{i} - i = 1$*, or* $\hat{j} - j = -1$*.*

An example illustrating Definition 2.8 is shown in Figure 1. It is straightforward to check that if $\lambda \succ \hat{\lambda}$, then $\lambda \ge \hat{\lambda}$ and further λ and $\hat{\lambda}$ are immediate neighbours in the dominance order.

Proof of Theorem 1.4. It suffices to consider the case $k = n-1$, as the case of general $k < n$ would follow from the former by induction. Further, due to the definition of the relation $\rho \ge \hat{\rho}$, it suffices to consider the case when both these measures are atoms, i.e. $sup(\rho) = \lambda$ and $sup(\hat{\rho}) = \hat{\lambda}$ with $\lambda \ge \hat{\lambda}$. Further, since the dominance order and stochastic dominance relation are transitive, it suffices to consider the case when λ and $\hat{\lambda}$ are immediate neighbors in the partial order, i.e. $\lambda \succ \hat{\lambda}$. Without loss of generality we assume that λ and $\hat{\lambda}$ differ by the move of the box $(i,j) \subset \lambda$ into the position $(\hat{i},\hat{j}) \subset \hat{\lambda}$ such that $\hat{i} - i = 1$.

In the latter case $\pi^n_{n-1}(\rho)$ assigns the mass

$$\frac{\dim(\mu)}{\dim(\lambda)} \tag{13}$$

to each diagram $\mu \in \mathbb{Y}_{n-1}$, such that $\mu \nearrow \lambda$. Similarly, $\pi^n_{n-1}(\hat{\rho})$ assigns the mass

$$\frac{\dim(\hat{\mu})}{\dim(\hat{\lambda})} \tag{14}$$

to each diagram $\hat{\mu} \in \mathbb{Y}_{n-1}$, such that $\hat{\mu} \nearrow \hat{\lambda}$. Subdivide all $\mu \in \mathbb{Y}_{n-1}$, such that $\mu \nearrow \lambda$ into three sets

$$A^{\uparrow}_\lambda = \{\mu \in \mathbb{Y}_{n-1} \mid \lambda \setminus \mu = (r,c),\ r < i\}, \quad A^{\downarrow}_\lambda = \{\mu \in \mathbb{Y}_{n-1} \mid \lambda \setminus \mu = (r,c),\ r > \hat{i}\},$$

$$A_\lambda^{=} = \{\mu \in \mathbb{Y}_{n-1} \mid \lambda \setminus \mu = (r,c),\ i \le r \le \hat{i}\}.$$

Now for $\mu \in A_\lambda^{\uparrow} \cup A_\lambda^{\downarrow}$ set $\hat{\mu} \prec \mu$ to be the Young diagram obtained by moving the box (i,j) into the position $(\hat{i},\hat{j})$. Now the following three observations imply the stochastic dominance $\pi_{n-1}^n \rho \ge \pi_{n-1}^n \hat{\rho}$:

- All the Young diagrams from $A_\lambda^{\uparrow} \cup A_\lambda^{=} \cup A_\lambda^{\downarrow}$ are linearly ordered (with respect to the dominance order) by r, which is the row number of the box being removed from λ. The same is true for $A_{\hat{\lambda}}^{\uparrow} \cup A_{\hat{\lambda}}^{=} \cup A_{\hat{\lambda}}^{\downarrow}$.

- Each Young diagram from $A_\lambda^{\downarrow} \cup A_\lambda^{=}$ dominates each Young diagram from $A_{\hat{\lambda}}^{\uparrow} \cup A_{\hat{\lambda}}^{=}$.

- Due to Corollary 2.6 and formulas (13), (14), for each $\mu \in A_\lambda^{\uparrow}$ we have $(\pi_{n-1}^n \rho)(\mu) \le (\pi_{n-1}^n \hat{\rho})(\hat{\mu})$ and for each $\mu \in A_\lambda^{\downarrow}$ we have $(\pi_{n-1}^n \rho)(\mu) \ge (\pi_{n-1}^n \hat{\rho})(\hat{\mu})$. $\square$

3 The Law of Large Numbers for the Young graph

The second ingredient of our proof of the Thoma theorem (Theorem 1.2) is the Law of Large Numbers for the measures appearing in its formulation.

Theorem 3.1 (The law of large numbers, [VK1], [KOO], [Bu], [Me])**.** *Choose two strictly decreasing finite sequences $\alpha_1 > \alpha_2 > \dots > \alpha_a > 0$, $\beta_1 > \beta_2 > \dots > \beta_b > 0$ such that $\sum_{i=1}^a \alpha_i + \sum_{i=1}^b \beta_i = 1$.*

For $n = 1, 2, \dots$ let $\lambda(n) \in \mathbb{Y}_n$ be a random Young diagram distributed according to the probability measure

$$M_n^{(\alpha,\beta)}(\lambda) = \dim(\lambda) s_\lambda(\alpha,\beta).$$

Then for each $i = 1, \dots, a$ and each $j = 1, \dots, b$ we have (in probability)

$$\lim_{n\to\infty} \frac{\lambda_i(n)}{n} = \alpha_i, \quad \lim_{n\to\infty} \frac{\lambda'_j(n)}{n} = \beta_j.$$

Remark 3.2. In fact, an analogue of Theorem 3.1 holds for all extreme measures of Theorem 1.2, see [VK1], [KOO], [Bu], [Me]. However, the present weaker form is enough for our purposes.

There are at least four different approaches in the literature to the proof of Theorem 3.1:

- The proof of the Thoma theorem in [VK1], [KOO] based on the relation of the dimensions in Young graph to the *shifted* Schur functions, as a byproduct implies Theorem 3.1. Note that we would like to avoid using this approach here, since our aim is to produce an independent proof of Thoma theorem.

- Vershik and Kerov in [VK2] showed how the random Young diagrams $\lambda(n)$ can be sampled using (a modification of) the classical Robinson–Schensted correspondence, whose input is a sequence of n i.i.d. discrete random variables. This observation allows to deduce Theorem 3.1 from the conventional Law of Large Numbers for sequences of independent random variables. For the details we refer to [Bu], where, in fact, a stronger Cental Limit Theorem was proved using this approach.

- Kerov explained in [K1] (see also [IO]) how certain observables of random Young diagrams $\lambda(n)$ can be computed using the algebra of shifted–symmetric functions. The resulting formulas turn out to be well-suited for the asymptotics analysis along the lines of Theorem 3.1, which was done in [Me]. In fact, [Me] also proves a stronger Central Limit Theorem.

- Following the approach of [J], [BOO], [Ok2] one proves that the *poissonization* of measures M_n can be described via a *determinantal point process*, with an explicit contour integral expression for the kernel. Asymptotic analysis of this kernel via steepest descent gives Theorem 3.1.

Each of the above four methods for proving Theorem 3.1 relies on a certain very nontrivial (but known) technique, which is the algebra of shifted–symmetric functions for the first and third approaches, the Robinson–Schensted correspondence for the second approach and determinantal point processes / Schur measures for the forth one. Given the knowledge of this technique the proof of Theorem 3.1 becomes relatively simple.

We now give a sketch of the second "combinatorial" proof of Theorem 3.1, which is based on the Robinson–Schensted correspondence.

Sketch of the proof of Theorem 3.1. Let us consider an alphabet $\mathcal{T} = \mathcal{T}^+ \cup \mathcal{T}^-$, where $\mathcal{T}^+ = \{t_1^+, \dots, t_a^+\}$ and $\mathcal{T}^- = \{t_1^-, \dots, t_b^-\}$. Let us fix a linear order on $\mathcal{T}$; its exact choice is irrelevant, so e.g. one can assume that

$$t_b^- < t_{b-1}^- < \cdots < t_1^- < t_1^+ < \cdots < t_q^+.$$

For $x, y \in \mathcal{T}$ we write $x \lhd y$ if either $x < y$, or $x = y \in \mathcal{T}^+$. We write $x \rhd y$ if either $x > y$ or $x = y \in \mathcal{T}^-$. We call a word $x_1 \dots x_n \in \mathcal{A}^n$ *increasing* if $x_1 \lhd x_2 \lhd \cdots \lhd x_n$, and *decreasing* if $x_1 \rhd x_2 \rhd \cdots \rhd x_n$. For a word w let us denote by $r_s(w)$ the maximal cardinality of the union of s disjoint increasing subsequences of the word w, and by $c_s(w)$ the maximal cardinality of the union of s disjoint decreasing subsequences.

Now let us define the probability measure $\eta^{(\alpha,\beta)}$ on $\mathcal{T}$ such that $\eta^{(\alpha,\beta)}(a_i) = \alpha_i$ and $\eta^{(\alpha,\beta)}(b_j) = \beta_j$. Let $w(n)$, $n = 1, 2 \dots$ be a random element of $\mathcal{T}^n$ distributed according to the product measure $(\eta^{(\alpha,\beta)})^{\otimes n}$. Vershik-Kerov [VK2] relying on a generalization of Robinson–Schensted correspondence (see also [BR]) proved that the following equality in distribution holds jointly for all $s = 1, 2, \dots$

$$\lambda_1(n) + \cdots + \lambda_s(n) \overset{d}{=} r_s(w(n)), \qquad \lambda'_1(n) + \cdots + \lambda'_s(n) \overset{d}{=} c_s(w(n)). \tag{15}$$

The identity (15) reduces Theorem 3.1 to the Law of Large Numbers as $n \to \infty$ for $r_s(w(n))$ and $c_s(w(n))$, $s = 1, 2, \dots$. The latter is rather transparent. Indeed, it is intuitively clear that the length of the longest increasing subsequence in the word $w(n)$ should be (up to a small error) equal to the length of the subsequence of all letters t_1^+ in $w(n)$, and the last length is approximately $\alpha_1 \cdot n$ due to the classical Law of Large Numbers for independent random variables. Further, the main contribution to $r_s(w(n))$ comes when each subsequence contains only one letter from our alphabet, and thus $r_s(w(n)) \approx (\alpha_1 + \cdots + \alpha_s) \cdot n$ for $1 \le s \le a$. Similarly, $c_s(w(n)) \approx (\beta_1 + \ldots, \beta_s) \cdot n$ for $1 \le s \le b$. A formal proof based on this argument is given in [Bu, Theorem 2], see also [Me, Section 6] and [S, Theorem 6.4]. $\square$

4 Proof of Thoma theorem

We start by explaining informally the main idea behind the proof of Theorem 1.2.

For any $\lambda \in \mathbb{Y}_k$ we define a probability measure ρ_λ on $\mathbb{Y}_k$ to be an atom with support $\sup(\rho_\lambda) = \lambda$. We start the proof from an abstract convex analysis statement (Proposition 4.1) that any extreme coherent system $\{M_n\}$ can be approximated by systems of the form $\pi_n^k(\rho_{\lambda(k)})$ for a sequence $\lambda(k) \in \mathbb{Y}_k$, $k = 1, 2, \ldots$. We further use the Law of Large Numbers to show in Lemma 4.4 that when k is large enough and after dropping out a tiny mass ε, the measure $\rho_{\lambda(k)}$ can be clutched between two measures $M_k^{(\alpha^-,\beta^-)}$ and $M_k^{(\alpha^+,\beta^+)}$. Moreover, they can be chosen so that the distance between (α^-, β^-) and (α^+, β^+) is small. Now Theorem 1.4 implies that $\pi_n^k(\rho_{\lambda(k)})$ is clutched between $M_n^{(\alpha^-,\beta^-)}$ and $M_n^{(\alpha^+,\beta^+)}$. At this point we conclude that any coherent system $\{M_n\}$ can be well-approximated by the coherent systems of the form $\{M_n^{(\alpha,\beta)}\}$, $(\alpha, \beta) \in \Omega$. Therefore, the closedness of the latter set of coherent systems implies Theorem 1.2.

The formal proof of Theorem 1.2 is given at the end of this section after we present a series of auxiliary statements.

Proposition 4.1. *Let $\{M_n\}_{n=1}^\infty$ be an extreme coherent system of measures. Then there exists a (deterministic) sequence of Young diagrams $\lambda(k) \in \mathbb{Y}_k$, $k = 1, 2, \ldots$ such that*

$$M_n = \lim_{k\to\infty} \pi_n^k(\rho_{\lambda(k)}), \quad n = 1, 2, \ldots. \tag{16}$$

Proof. This is a particular case of a very general convex analysis statement, which was reproved many times in different contexts. Its first appearance in the asymptotic representation theory dates back to [V], since then it is known as "ergodic method". The complete proofs of the statements generalizing Proposition 4.1 can be found in [OO, Section 6] or [DF, Theorem 1.1]. $\square$

Recall that for two measures ρ, $\hat\rho$ on a finite set A, their total variation distance is defined through

$$\mathbf{d}_{\mathrm{var}}(\rho, \hat\rho) = \frac{1}{2} \sum_{a \in A} |\rho(a) - \hat\rho(a)|.$$

We also define the L_∞ distance between two pairs of sequences $(\alpha, \beta) = (\alpha_1 \ge \alpha_2 \ge \ldots, \beta_1 \ge \beta_2 \ge \ldots)$, $(\hat\alpha, \hat\beta) = (\hat\alpha_1 \ge \hat\alpha_2 \ge \ldots, \hat\beta_1 \ge \hat\beta_2 \ge \ldots)$ through

$$\mathbf{d}_\infty((\alpha, \beta), (\hat\alpha, \hat\beta)) = \max\left(\sup_i |\alpha_i - \hat\alpha_i|, \sup_i |\beta_i - \hat\beta_i| \right).$$

The following two lemmas explain that the metrics $\mathbf{d}_{\mathrm{var}}$ on probability measures on $\mathbb{Y}_n$ and $\mathbf{d}_\infty$ on Ω are compatible.

Lemma 4.2. *For any $n = 1, 2, \dots$ we have*

$$\lim_{\varepsilon \to 0} \sup_{\substack{(\alpha,\beta),(\hat\alpha,\hat\beta)\in\Omega: \\ \mathbf{d}_\infty((\alpha,\beta),(\hat\alpha,\hat\beta))\le\varepsilon}} \mathbf{d}_{\mathrm{var}}(M_n^{(\alpha,\beta)}, M_n^{(\hat\alpha,\hat\beta)}) = 0. \tag{17}$$

Proof. Note that $\mathbb{Y}_n$ is a finite, therefore it suffices to prove (17) with $\mathbf{d}_{\mathrm{var}}$ replaced by $|M_n^{(\alpha,\beta)}(\lambda) - M_n^{(\hat\alpha,\hat\beta)}(\lambda)|$ for arbitrary $\lambda \in \mathbb{Y}_n$. Moreover, due to the definition (1), it suffices to study $|s_\lambda(\alpha,\beta) - s_\lambda(\hat\alpha,\hat\beta)|$. To analyze this difference recall that the Schur function s_λ is a polynomial in power sums $p_1, \dots, p_n$, which generate the algebra of symmetric functions. We conclude that (17) is equivalent to

$$\lim_{\varepsilon \to 0} \sup_{\substack{(\alpha,\beta),(\hat\alpha,\hat\beta)\in\Omega: \\ \mathbf{d}_\infty((\alpha,\beta),(\hat\alpha,\hat\beta))\le\varepsilon}} |p_n(\alpha,\beta) - p_n(\hat\alpha,\hat\beta)| = 0, \quad n = 1, 2, \dots. \tag{18}$$

To prove (18) we recall the definition (2) and first conclude that

$$|p_1(\alpha,\beta) - p_1(\hat\alpha,\hat\beta)| = |1 - 1| = 0.$$

Further, for $n > 1$ we have

$$\begin{aligned} |p_n(\alpha,\beta) - p_n(\hat\alpha,\hat\beta)| &\le \sum_{i=1}^{\infty} |\alpha_i - \hat\alpha_i| \big((\alpha_i)^{n-1} + (\alpha_i)^{n-2}(\hat\alpha_i)^1 + \dots + (\hat\alpha_i)^{n-1} \big) \\ &\quad + \sum_{i=1}^{\infty} |\beta_i - \hat\beta_i| \big((\beta_i)^{n-1} + (\beta_i)^{n-2}(\hat\beta_i)^1 + \dots + (\hat\beta_i)^{n-1} \big) \\ &\le \mathbf{d}_\infty((\alpha,\beta),(\hat\alpha,\hat\beta)) \cdot n \sum_{i=1}^{n} \Big[(\alpha_i)^{n-1} + (\hat\alpha_i)^{n-1} + (\beta_i)^{n-1} + (\hat\beta_i)^{n-1} \Big] \\ &\le 4n \cdot \mathbf{d}_\infty((\alpha,\beta),(\hat\alpha,\hat\beta)), \end{aligned} \tag{19}$$

which immediately implies (18). □

Lemma 4.3. *Let $(\alpha(k), \beta(k))$, $k = 1, 2, \dots$ be pairs of sequences. Suppose that for each $n = 1, 2, \dots$ the measures $M_n^{(\alpha(k),\beta(k))}$ converge in the sense of $\mathbf{d}_{\mathrm{var}}$ to a measure M_n. Then there exists a pair of sequences (α, β) such that $M_n = M_n^{(\alpha,\beta)}$ for all n.*

Proof. We first claim that Ω is a compact set in the topology defined by $\mathbf{d}_\infty$. Indeed, this topology on Ω is equivalent to the topology of pointwise convergence. For the latter topology Ω is compact, since it is a closed subset of the compact set $[0,1]^\infty$. Now we define (α,β) as a limiting point of the sequence of pairs $(\alpha(k),\beta(k))$, $k = 1, 2, \dots$. Using Lemma 4.2 we conclude that $M_n = M_n^{(\alpha,\beta)}$ for all n. □

The next lemma is the key point of our proof of Theorem 1.2.

Lemma 4.4. *Take a sequence of integers $0 < k(1) < k(2) < \dots$ and let $\lambda(n) \in \mathbb{Y}_{k(n)}$, $n = 1, 2, \dots$ be a sequence of Young diagrams such that the following limits exist for each $i = 1, 2, \dots$*

$$\lim_{n\to\infty} \frac{\lambda_i(n)}{k(n)} = \alpha_i, \quad \lim_{n\to\infty} \frac{\lambda_i'(n)}{k(n)} = \beta_i.$$

Then for every $\varepsilon > 0$ and every $N \in \mathbb{N}$ there exists $n > N$, two measures ρ_n^+, ρ_n^- on $\mathbb{Y}_{k(n)}$ and two pairs of sequences $(\alpha^+, \beta^+), (\alpha^-, \beta^-) \in \Omega$, such that

1. $\mathbf{d}_{\mathrm{var}}\left(\rho_n^-, M_{k(n)}^{(\alpha^-,\beta^-)}\right) < \varepsilon$ *and* $\mathbf{d}_{\mathrm{var}}\left(\rho_n^+, M_{k(n)}^{(\alpha^+,\beta^+)}\right) < \varepsilon$,
2. $\mathbf{d}_\infty((\alpha^-, \beta^-), (\alpha^+, \beta^+)) < \varepsilon$,
3. $\rho_n^- \le \rho_{\lambda(n)} \le \rho_n^+$ *in the sense of stochastic dominance.*

In words, Lemma 4.4 says that the delta–measure on a Young diagram of a large level $\mathbb{Y}_k$ (after dropping a tiny mass ε) can be always clutched between two measures $M_k^{(\alpha^-,\beta^-)}$ and $M_k^{(\alpha^+,\beta^+)}$. Moreover, they can be chosen so that the distance between (α^-, β^-) and (α^+, β^+) is small. The proof relies on the Law of Large Numbers for the measures $M_k^{(\alpha,\beta)}$.

Proof of Lemma 4.4. Take $L_\alpha, L_\beta > 0$ such that $\alpha_{L_\alpha} < \varepsilon/2$ and $\beta_{L_\beta} < \varepsilon/2$, but $\alpha_i \ge \varepsilon/2$ for all $i < L_\alpha$ and $\beta_j \ge \varepsilon/2$ for all $j < L_\beta$. Further choose $V > 2$, such that $\alpha_{L_\alpha} < \varepsilon/2 - \varepsilon/V$ and $\beta_{L_\beta} < \varepsilon/2 - \varepsilon/V$. We will now define the pair of sequences (α^+, β^+) as follows.

$$\alpha_i^+ = \alpha_i + \frac{\varepsilon}{V \cdot 2^i}, \quad i = 2, \dots, L_\alpha, \qquad \beta_j^+ = \beta_j - \frac{\varepsilon}{V \cdot 2^{L_\beta + 1 - j}}, \quad j = 1, \dots, L_\beta.$$

For $j > L_\beta$ we set $\beta_i^+ = 0$. For $i = L_\alpha + 1, \dots, R$ we set $\alpha_i^+ = \varepsilon/2 - \varepsilon/V + \frac{\varepsilon}{V2^i}$ where R is the minimum integer such that

$$S(R) := \left(\alpha_1 + \frac{\varepsilon}{2V}\right) + \sum_{i=2}^{R+1} \alpha_i^+ + \sum_{j=1}^{L_\beta} \beta_j^+ > 1.$$

Finally, set $\alpha_1^+ = \alpha_1 + \frac{\varepsilon}{2V} + (1 - S(R-1))$ and $\alpha_i = 0$ for $i > R$.

Note that the resulting (α^+, β^+) satisfies the assumptions of Theorem 3.1. Combining this theorem with the definition of numbers α_i, β_i, we conclude the existence of N_1 such that for all $n > N_1$ the diagram $\lambda(n) \in \mathbb{Y}_{k(n)}$ is dominated by $M_{k(n)}^{(\alpha^+,\beta^+)}$–random Young diagram $\mu(n)$ with probability greater than $(1 - \varepsilon)$. Thus, if we define $\rho^+(n)$ on $\mathbb{Y}_{k(n)}$ through the identity

$$\rho_n^+(\mu) = \begin{cases} M_{k(n)}^{(\alpha^+,\beta^+)}(\mu), & \mu > \lambda(n), \\ 1 - \sum_{\nu > \lambda(n)} M_{k(n)}^{(\alpha^+,\beta^+)}(\nu), & \mu = \lambda(n), \\ 0, & \text{otherwise.}, \end{cases}$$

then both $\mathbf{d}_{\mathrm{var}}\left(\rho^+, M_{k(n)}^{(\alpha^+,\beta^+)}\right) < \varepsilon$ and $\rho_{\lambda(n)} \le \rho_n^+$ hold.

Arguing similarly but with the roles of α's and β's switched, we define (α^-, β^-) and $\rho^-(n)$. It remains to note that

$$\mathbf{d}_\infty((\alpha^-,\beta^-),(\alpha^+,\beta^+)) \le \mathbf{d}_\infty((\alpha^-,\beta^-),(\alpha,\beta)) + \mathbf{d}_\infty((\alpha,\beta),(\alpha^+,\beta^+)) < \varepsilon/2+\varepsilon/2 = \varepsilon. \quad \square$$

Proof of Theorem 1.2. Let $\{M_r\}_{r=1}^\infty$ be an extreme coherent system of measures and let $\lambda(k) \in \mathbb{Y}_k$, $k = 1, 2, \ldots$ be a corresponding sequence of Young diagrams as in Proposition 4.1. Since for all $i = 1, 2, \ldots,$ we have $0 \le \lambda_i(k)/k \le 1$ and $0 \le \lambda_i'(k) \le 1$, passing to a subseqence $k(n)$, $n = 1, 2, \ldots$ we can assume that the following limits exist

$$\lim_{n\to\infty} \frac{\lambda_i(k(n))}{k(n)} = \alpha_i, \quad \lim_{n\to\infty} \frac{\lambda_i'(k(n))}{k(n)} = \beta_i.$$

Now we choose $\varepsilon(n) = 1/n$. Passing, if necessary, to another subsequence (which we will denote by the same $k(n)$ to avoid complicating the notations) and using Lemma 4.4, we conclude that there exist $(\alpha^-(n), \beta^-(n)), (\alpha^+(n), \beta^+(n)) \in \Omega$ and measures $\rho^+(n), \rho^-(n)$ on $\mathbb{Y}_{k(n)}$ such that

1. $\mathbf{d}_{\text{var}}\left(\rho^-(n), M_{k(n)}^{(\alpha^-(n),\beta^-(n))}\right) < \frac{1}{n}$ and $\mathbf{d}_{\text{var}}\left(\rho^+(n), M_{k(n)}^{(\alpha^+(n),\beta^+(n))}\right) < \frac{1}{n}$,
2. $\mathbf{d}_\infty\big((\alpha^-(n),\beta^-(n)),\, (\alpha^+(n),\beta^+(n))\big) < \frac{1}{n}$,
3. $\rho^-(n) \le \rho_{\lambda(k(n))} \le \rho^+(n)$ in the sense of stochastic dominance.

Now choose any $r = 1, 2, \ldots.$ We aim to prove that $M_r = \lim_{n\to\infty} M_r^{(\alpha^-(n),\beta^-(n))}$ in the sense of $\mathbf{d}_{\text{var}}$. For that note that since each map π_k^m is a contraction in $\mathbf{d}_{\text{var}}$ distance, Lemma 4.2 implies as $n \to \infty$

$$\begin{aligned} \mathbf{d}_{\text{var}}(\pi_r^{k(n)}\rho^-(n), \pi_r^{k(n)}\rho^+(n)) &\le \mathbf{d}_{\text{var}}(\pi_r^{k(n)}\rho^-(n), M_r^{(\alpha^-(n),\beta^-(n))}) \\ &+ \mathbf{d}_{\text{var}}\left(M_r^{(\alpha^-(n),\beta^-(n))}, M_r^{(\alpha^+(n),\beta^+(n))}\right) + \mathbf{d}_{\text{var}}(M_r^{(\alpha^+(n),\beta^+(n))}, \pi_r^{k(n)}\rho^+(n)) \\ &\le \frac{2}{n} + \mathbf{d}_{\text{var}}\left(M_r^{(\alpha^-(n),\beta^-(n))}, M_r^{(\alpha^+(n),\beta^+(n))}\right) \to 0. \end{aligned} \quad (20)$$

We claim that the last inequality implies that

$$\mathbf{d}_{\text{var}}\left(\pi_r^{k(n)}\rho^-(n), \pi_r^{k(n)}\rho_{\lambda(k(n))}\right) \to 0. \quad (21)$$

Indeed, by Theorem 1.4

$$\pi_r^{k(n)}\rho^-(n) \le \pi_r^{k(n)}\rho_{\lambda(k(n))} \le \pi_r^{k(n)}\rho^+(n).$$

Thus, for any upper[2] set $U \subset \mathbb{Y}_r$ we have

$$\pi_r^{k(n)}\rho^-(n)(U) \le \pi_r^{k(n)}\rho_{\lambda(k(n))}(U) \le \pi_r^{k(n)}\rho^+(n)(U).$$

[2]By the definition an upper set U in a partially ordered set A satisfies the property that if $x \in U$ and for some $y \in A$ we have $x < y$, then also $y \in U$.

Therefore, as $n \to \infty$

$$|\pi_r^{k(n)} \rho^-(n)(U) - \pi_r^{k(n)} \rho_{\lambda(k(n))}(U)| \le |\pi_r^{k(n)} \rho^-(n)(U) - \pi_r^{k(n)} \rho^+(n)(U)| \\ \le \mathbf{d}_{\mathrm{var}}(\pi_r^{k(n)} \rho^-(n), \pi_r^{k(n)} \rho^+(n)) \to 0. \quad (22)$$

Note that for any $\lambda \in \mathbb{Y}_r$ both $\{\mu \in \mathbb{Y}_r : \mu \ge \lambda\}$ and $\{\mu \in \mathbb{Y}_r : \mu > \lambda\}$ are upper sets, whose difference is $\{\lambda\}$. Therefore, (22) implies (21). Now combining (21) with (16) and with inequality $\mathbf{d}_{\mathrm{var}}(\pi_r^{k(n)} \rho^-(n), M_r^{(\alpha^-(n),\beta^-(n))}) \le 1/n$, we prove that

$$M_r = \lim_{n\to\infty} M_r^{(\alpha^-(n),\beta^-(n))}.$$

Now it remains to apply Lemma 4.3. □

References

[BDJ] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, Journal of the American Mathematical Society, 12, no. 4, 1119– 1178, (1999). arXiv:math/9810105.

[BR] A. Berele and A. Regev, Hook Young diagrams with applications to combinatorics and representations of Lie superalgebras. Advances in Mathematics, 64 (1987), 118–175.

[B] A. M. Borodin, The law of large numbers and the central limit theorem for the jordan normal form of large triangular matrices over a finite field, Journal of Mathematical Sciences (New York), 1999, 96:5, 3455–3471

[BC] A. Borodin, I. Corwin, Macdonald processes, Probability Theory and Related Fields, 158, no. 1-2 (2014), 225–400, arXiv:1111.4408.

[BCGS] A. Borodin, I. Corwin, V. Gorin, S. Shakirov, Observables of Macdonald processes, to appear in Transactions of American Mathematical Society, arXiv:1306.0659.

[BOO] A. Borodin, A. Okounkov and G. Olshanski, Asymptotics of Plancherel measures for symmetric groups, Journal of American Mathematical Society, 13 (2000) 491–515. arXiv:math/9905032.

[Bu] A. Bufetov, The central limit theorem for extremal characters of the infinite symmetric group, Functional Analysis and Its Applications, 46,no. 2 (2012), 83–93, arXiv:1105.1519

[BP] A. Bufetov, L. Petrov, Law of Large Numbers for Infinite Random Matrices over a Finite Field, arXiv:1402.1772.

[DF] P. Diaconis, D. Freedman, Partial Exchangeability and Sufficiency. Proc. Indian Stat. Inst. Golden Jubilee Int'l Conf. Stat.: Applications and New Directions, J. K. Ghosh and J. Roy (eds.), Indian Statistical Institute, Calcutta (1984), pp. 205-236.

[Fu] J. Fulman, Random matrix theory over finite fields, Bulletin of American Mathematical Society 39 (2002), 51-85.

[GKV] V. Gorin, S. Kerov, A. Vershik, Finite traces and representations of the group of infinite matrices over a finite field, *Advances in Mathematics*, 254 (2014), 331–395. arXiv:1209.4945.

[IO] V. Ivanov, G. Olshanski, Kerov's central limit theorem for the Plancherel measure on Young diagrams, In: S.Fomin, editor. Symmetric Functions 2001: Surveys of Developments and Perspectives (NATO Science Series II. Mathematics, Physics and Chemistry. Vol.74), Kluwer, 2002, 93–151, arXiv:math/0304010

[J] K. Johansson, Discrete orthogonal polynomial ensembles and the Plancherel measure, Annals of Mathematics (2) 153 (2001), no. 2, 259–296. arXiv:math/9906120.

[K1] S. Kerov, Gaussian limit for the Plancherel measure of the symmetric group, Comptes Rendus Acad. Sci. Paris, Serie I 316 (1993), 303–308

[K2] S. Kerov: Asymptotic Representation Theory of the Symmetric Group and its Applications in Analysis, Amer. Math. Soc., Providence, RI, 2003.

[KOO] S. Kerov, A. Okounkov, G. Olshanski, The boundary of Young graph with Jack edge multiplicities, Interntional Mathematics Research Notices, no. 4 (1998), 173–199. arXiv:q-alg/9703037

[Kin] J. F. C. Kingman, Random partitions in population genetics, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 361 (1978), 1–20.

[Kir] A. A. Kirillov, Variations on the triangular theme, Dynkin Seminar on Lie Groups, Advances in Mathematical Sciences, Series 2, 169 (1995), AMS.

[LP] T. Lam, P. Pylyavskyy, Cell transfer and monomial positivity, Journal of Algebraic Combinatorics, 26, no. 2 (2007), 209–224, arXiv:math/0505273

[LPP] T. Lam, A. Postnikov, P. Pylyavskyy, Schur positivity and Schur log-concavity, American Journal of Mathematics, 129, no. 6 (2007), 1611–1622. arXiv:math.CO/0502446

[Ma] I. G. Macdonald, Symmetric functions and Hall polynomials, Second Edition. The Clarendon Press, Oxford University Press, New York, 1995

[Me] P. L. Méliot. A central limit theorem for the characters of the infinite symmetric group and of the infinite Hecke algebra, arXiv:1105.0091

[Ok1] A. Okounkov, On the representations of the infinite symmetric group, Zapiski Nauchnyh Seminarov POMI, 240 (1997), 167–230, arXiv:math/9803037.

[Ok1] A. Okounkov, Random Matrices and Random Permutations, International Mathematics Research Notices (2000) 2000 (20), arXiv:math/9903176

[Ok2] A. Okounkov. Infinite wedge and random partitions. Selecta Mathematica 7 (2001), 57–81. arXiv:math/9907127.

[OO] A. Okounkov, G. Olshansky, Asymptotics of Jack Polynomials as the Number of Variables Goes to Infinity, International Mathematics Research Notices 13 (1998), pp. 641–682.

[S] P. Sniady, Robinson-Schensted-Knuth algorithm, jeu de taquin and Kerov-Vershik measures on infinite tableaux, SIAM Journal of Discrete Mathematics 28 , no. 2 (2014), 598–630, arXiv:1307.5645.

[T] E. Thoma: Die unzerlegbaren, positiv-definiten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe, Mathematische Zeitschrift, 85 (1964), 40–61.

[V] A. M. Vershik, Description of invariant measures for the actions of some infinite–dimensional groups, Soviet Mathematics Doklady 15 (1974), 1396–1400.

[V2] A. M. Vershik, The problem of describing central measures on the path spaces of graded graphs, arXiv:1408.3291

[VK1] A. M. Vershik, S. V. Kerov, Asymptotic character theory of the symmetric group, Functional Analysis and its Applications, 15 (1981), 246–255.

[VK2] A. M. Vershik, S. V. Kerov, The characters of the infinite symmetric group and probability properties of the Robinson–Schensted–Knuth algorithm, SIAM Journal on Algebraic Discrete Methods, 7, no. 1 (1986), 116–124.